Implementation Guidance to the USEPA Region 2 Draft Interim Policy on Identifying EJ Areas

June, 1999

Part II

Prepared by the Interim Policy Subgroup of the Region 2 Environmental Justice Work Group

Implementation Guidance to the
USEPA Region 2 Draft Interim Policy on Identifying EJ Areas
Table of Contents

Preface

USEPA Region 2 has developed a "Draft Interim Policy on Identifying EJ Areas,"June, 1997 (Revised, June 1999). That document addresses one of the most important aspects of any governmental Environmental Justice (EJ) programs, identifying communities or areas that may be potential or actual EJ communities or that may pose EJ concerns. It sets forth the Region 2 policy, defining terms, summarizing the steps that are to be taken in preparing for an EJ assessment, and specifying the Decision-criteria that are to be used in making the actual assessments. Once an area is assessed to be an EJ area, subsequent Agency actions would be in accordance with established laws, regulations, and policies.

This document is a companion to that Interim Policy. It repeats definitions and policy statements, and provides considerably more detail on how to carry out the process. It provides all of the terms, definitions, methodologies, and guidance needed for identifying EJ areas in Region 2.

This document is divided into three sub-parts. The first sub-part - The Policy and the Process - repeats and expands somewhat the information contained in the Interim Policy. It is designed for people who are interested in the process but will not be collecting the data and performing the actual analyses. It clearly states the policy and lists the steps that must be followed in order to assess whether an area is an EJ Area. The second sub-part - Implementation Details - describes the nuts and bolts of how to proceed with each step listed in Part I. These details involve a number of specialized technical operations and require a considerable level of experience. While the steps are clear, the final assessment for an area may be far from obvious, due to the large amount of information that is brought to bear on the issue, and the inherent uncertainty in much of that information. Sub-Part III contains the Appendices, which serve to support and explain elements of both Sub-Parts I and II.

It is recommended that members of the public or managers interested in the consequences of the process utilize the Interim Policy itself, or pay close attention to Sub-Part I of this document, but only deal with Sub-Part II as needed. On the other hand, people who will be performing the process for recommending the EJ status of an area will need to pay close attention to both Sub-Parts I and II of this document, using Sub-Part II essentially as a standard operating procedure for the process.
The statements in this document are intended solely as guidance. This document is not intended, nor can it be relied upon, to create any rights enforceable by any party in litigation with the United States. EPA Region 2 may decide to follow the guidance provided in this document, or to act at variance with the guidance based on its analysis of the specific facts presented. This Draft Interim Guidance may be revised without public notice to reflect changes in EPA's approach to implementing Executive Order 12898.

Sub-Part I - The Policy and the Process

1. Draft Interim Policy on Identifying EJ Areas

a. Introduction

The Region 2 Draft Interim Policy on Identifying Environmental Justice Areas, June 1997 (Revised , June 1999), (Draft Interim Policy) defines and describes the process to be utilized in U.S. EPA Region 2 (Region 2) for assessing whether a specific area is subject to the Agency's Environmental Justice (EJ) Program. This companion document expands on the Draft Interim Policy, explains in detail how the Draft Interim Policy is to be implemented, and contains the necessary supporting information. Part I describes the need for the process, lists all of the steps in the process, defines the Decision-criteria that must be satisfied if an area is to be considered an EJ Area, and offers guidance for applying those Decision-criteria. Part II contains all of the detailed definitions and explanations necessary for technically trained and experienced individuals to implement the Interim Policy in a given situation. Part III contains Appendices that explain and support the first two parts. The Glossary at the end of Part II contains definitions for many of the terms used throughout this document.

b. Environmental Justice Terms and Definitions

The Interim Policy is an outgrowth of Executive Order 12898, issued on February 11, 1994, which directed each Federal agency to:

> make achieving environmental justice part of its mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental burdens of its programs, policies, and activities on minority populations and low-income populations in the United States and its territories and possessions, the District of Columbia, the Commonwealth of Puerto Rico and the Commonwealth of the Mariana Islands.

The Interim Policy uses the term "minority" rather than "people of color" in order to be consistent with the Executive Order, but the Region is mindful and supportive of many communities' desire to be identified as "people of color." The Interim Policy uses the term "American Indian" in referring to all indigenous populations within the Region, regardless of their affiliation with a federally-recognized Tribe, but recognizes various terminology preferences and will strive to respect and utilize appropriate language on a case-by-case basis.

EPA Region 2 uses the terms "EJ Area" or "EJ Community" to describe a community that satisfies the terms of the Executive Order and the accompanying definitions. In addition, the terms related to'Adverse Environmental Burden' are frequently shortened to 'Burden' or 'Disproportionate Burden,' but the modifiers 'Adverse' and 'Environmental' are always implied.

c. The Process

The steps for identifying and screening EJ areas center on the comparison of three factors between the Community of Concern and one or more reference communities: their respective levels of minority representation, low-income representation, and environmental burden. Decision-criteria offer guidelines for determining whether the levels in the Community of Concern are significantly greater than the reference areas for

- minority **and/or** low-income, **and**
- environmental burden.

This Interim Policy specifies a procedure for making EJ assessments in a consistent and defensible way. This guidance offers a standard methodology for evaluating demographic burden data. It provides flexibility for how decision- makers weigh burden data given the unique situation of the COC. There are two approaches for identifying EJ areas:

- Screening analyses to identify potential EJ areas that warrant further study.
- Site-specific analyses to address EJ concerns.

EJ screening analyses are based on the consideration of demographic data, and do not focus on the assessment of disproportionate and adverse burden. Screening analyses address the demographic characteristics of geographical units within the study area, such as census blocks or block groups, municipalities or counties. Screening studies may utilize Geographic Information Systems (GIS), such as the Region 2 OPM Environmental Justice GIS, or an alternative approach such as the census centroid pull technique (Reminder, see the Glossary for definitions of less common terms, such as these). In either case, the focus of a screening analysis is on comparison of the demographic characteristics of discreet geographical communities of concern to a reference area that encompasses all of the communities of concern.

Site-specific-analyses will necessarily be more in-depth than screening analyses because a number of potentially difficult assessments must be made along the way. First, specific reference areas must be selected, their boundaries delineated, and their demographic data collected. Then, a site specific analysis requires a detailed analysis of the environmental burden in the Community of Concern and the reference communities in order to determine whether the burden is disproportionate and adverse in the Community of Concern.

The seven steps described below include as much specific guidance as possible, along with appropriate flexibility. All five steps are required for an identification of an EJ Area. However, only Steps 1 through 4 are required for a typical screening analysis, since environmental burden is not considered, and since an assessment of potential and not actual EJ areas is made. This compilation of the seven steps is intended as a checklist, to ensure that the process is implemented properly.

2. Region 2 Process for Screening and Identifying EJ Areas

a. Introduction

This Section outlines the seven major steps in the Region 2 process for screening and identifying EJ areas (Steps 1 - 7). As an aid to implementation, the explanations of the Steps are more detailed than in the Interim Policy itself. Section 3, which follows, contains the specific guidelines for applying the Decision-criteria required in Step 7.

b. Steps for Screening and Identifying EJ Areas

The seven steps described below include specific guidance for implementing the Interim Policy, along with appropriate flexibility. All seven steps are required for an identification of an EJ Community. However, only Steps 1 through 4, are required for a typical screening analysis, since environmental burden is not considered. This compilation of the seven steps is intended as a checklist, to ensure that the process is implemented properly.

The parenthetical number following each step or sub-step indicates the subsequent sections and subsections of this Implementation guidance document where the specific step is described in detail (for example, "(5)" shown at Step 1.i. means that Section 5 contains the details for performing that portion of Step 1).

This compilation of the seven steps is intended as a checklist.

Step 1: Define the Community of Concern

In order to perform an analysis for identifying an EJ area, the specific boundaries of the Community of Concern must be carefully defined and characterized.

i. Specify Geographic Boundaries (5.)
- Community-based definition
- Census-based definition
- Use of physical (spatial) criteria other than Census data via GIS
- Exposure-based boundaries (exposure modeling)

ii. Obtain Demographic Characteristics within the defined Geographic Boundaries (5.c.)
- Community-defined attributes
- Attributes based on Census data

Step 1 Results:

1. Clear geographic/boundary definition of the community of Concern.
2. Specific data describing the prevalent Minority and Low-Income demographic characteristics of the defined Community of Concern.

Step 2: Define the Reference Areas

Reference (comparison) areas must be defined to provide a context for interpretation

of data from the Community of Concern. For site-specific analyses, demographic characteristics of the reference areas should be similar to or in some way parallel to the characteristics of the Community of Concern. For screening analyses the reference areas are usually a large area, such as a state, that encompasses a number of possible Communities of Concern.

i. Specify Geographic Boundaries (5.)
   - Community-based selection process
   - Proximate data units (e.g., Tracts, zip codes): immediately adjacent units to the Community of Concern, a ring of units with a predetermined radius, etc.
   - Analogous data units in another area, similar in some respect to the Community of Concern, such as an area with a similar facility, a facility managed by the same organization, or an area subject to the same regulations
   - Proximate or adjacent communities
   - County, city, or other political jurisdiction in which the Community of Concern is located

ii. Obtain Demographic Characteristics of the defined reference areas (5.c.)

Step 2 Results:

1. Clear geographic/boundary definition of the reference areas.
2. Specific data describing the prevalent Minority and Low-Income demographic characteristics of the defined reference areas.

Step 3 & 4. Evaluate Minority and Income Data

Once the demographic data have been collected, it is necessary to evaluate them in detail, in order to put the information in a form that can be compared logically from one area to another.

i. Conduct the evaluation (including extraction and aggregation) of the demographic data in the Community of Concern. (6.)
ii. Conduct the evaluation (including extraction and aggregation) of the demographic data in the reference areas. (6.)

Step 4 Results:

1. Clear demographic statement for the Community of Concern.
2. Clear demographic statement for the reference areas.

Step 5. Develop the Environmental Load Profile

For communities that meet either demographic criterion, environmental load profiles must be developed for the Community of Concern and for the reference areas. Since methodologies to quantify actual burden are currently not available, the IP adopts the concept of developing

an environmental load profile for a community. The environmental load profile is comprised of different elements that when combined are indicative of burden. It is be based on salient characteristics that would serve as indicators of environmental burden and could provide a consistent basis for comparison.

i. Determine the environmental load profile in the Community of Concern.
ii. Determine the environmental load profile in the reference areas.

Step 3 Results:

1. Clear definition of the environmental load profile in the Community of Concern, using GIS tools.
2. Clear definition of the environmental load profile in the reference areas, using GIS tools.

The profiles will be developed using standardized methodologies with data sets that are available for the entire Region. These methodologies will often be GIS-based tools.

Step 5 Results:

1. Compilation of data for Community of Concern environmental load profile.
2. Compilation of data for reference areas environmental load profiles.

i. Assess the relative environmental burden in the Community of Concern.

Step 6. Evaluation of the Environmental Load Profile

Once the burden data for the environmental load profile have been collected, it is necessary to evaluate them in detail in order to put the information in a form that can be compared logically from one area to another.

Step 6. Results:

1. Assessment of whether the burden is adverse using the Decision-criteria in this Draft Interim Policy (4.d., 4.e.ii.).
2. Assessment of proportionality by comparing the environmental load profiles of the Community of Concern and reference communities (4.d.,4.e.ii.).

.

Step 7. Apply the Decision-Criteria

At this step, all of the data have been evaluated. The demographic and environmental burden results, from the environmental load profiles, for the Community of Concern have been compared with those from the reference areas. This comparison should be conducted according to the Decision-criteria for each EJ factor as follows:

i. Apply the Minority Decision-criterion to the Community of Concern and each

Reference Area (3.b.).

ii. Apply the Low-Income Decision-criterion to the Community of Concern and each Reference Area (3.c.).

iii. Apply the Environmental Load Profile Burden Decision-criteria to the Community of Concern and each Reference Area to assess if there is a disproportionality of burden (3.d.).

iv. Then, combine the results to produce a single document, including all of the appropriate supporting documentation and the assessment of whether:

- the Community of Concern has a greater minority and/or low-income population than the reference areas, and

- the Community of Concern suffers from a disproportionate and adverse environmental burden (3.a.).

In other words, assess whether the Community of Concern is an EJ Community.

Step 7. Results: 1. A documented assessment as to whether the Community of Concern satisfies the criteria to be considered an EJ community.

3. Designated Decision-criteria for EJ Factors

a. Introduction

This Section contains the appropriate Decision-criteria for comparing the Community of Concern and the reference areas that are required in Step 7 in Section 2. By applying the specified guidelines, Region 2 management will be able to determine whether the population in the Community of Concern is significantly higher in minority residents and/or low-income residents, and suffers from a significantly greater environmental burden, namely, that it is an EJ Area.

The policy requires that an EJ assessment include a justification for the decision. An underlying assumption to all of the Decision-criteria that follow is that the Region will use the data sources and techniques outlined here when they are available, including quantitative comparisons of the demographic characteristics of the Community of Concern and the reference communities.

For the minority and low-income EJ factors, the specific values that serve as break-points or cut-offs were based heavily upon a review of a large number of EJ studies for waste sites around the country (described in Appendix C). This review indicated clear distributions of the relative differences between the Community of Concern and the reference minority and low-income populations.

A clear majority of EJ researchers across the country identified populations as low-income or minority, when the percent differences were greater than the cut-offs specified in the guidelines that follow. As such, these Decision-criteria are conservative when compared with those

research results, and do not preclude the use of lower percent differences for identifying minority and/or low-income populations.

The cut-offs provided for the disproportionate and adverse environmental burden factor are based on the experience of Region 2 staff, since a database of quantified burdens in historical EJ assessments is not yet available. These cut-offs may serve only as a rough guide when non-quantifiable or unique burden is at issue. The Region may be able to develop more supportable cut-offs as more experience is gained with different types of environmental burdens.

The Decision-criteria specified below are expected to apply to most Region 2 EJ assessments. For the purpose of this Draft Interim Policy evaluation , the U.S. Census results will be used to assess if a Community of Concern meets the minority and low-income criteria established in this document. However, there will be certain situations in which the underlying assumptions are not appropriate. These special demographic and environmental considerations are extraordinary circumstances or exceptions to the prescribed methodology. A "special demographic" consideration, for example, may involve a Community of Concern that is indistinguishable from its neighbors or reference communities with respect to one of the demographic factors. "Special environmental" considerations, on the other hand, refer to unconventional exposure scenarios such as a Community of Concern living in close proximity to an airport or toll plaza . Appendix D addresses these issues in greater detail and provides additional information to facilitate the decision-making process. It is strongly recommended that the analyst routinely verify results to ensure that no special considerations exist. Failure to do so may have a significant bearing on whether the EJ criteria have, or have not, been satisfied.

b. Decision-criterion for the Minority Factor

The assessment of whether a community is minority will be based on a comparison of the Community of Concern with one or more reference communities, utilizing the same analytical methodology for each. For EJ purposes, a minority community is defined as one in which the percentage of minority residents is significantly greater than the comparable percentage in the reference community, after accounting for the accepted uncertainty in the data. The following guidelines are used for this assessment:

If the relative difference in the minority percentages between the Community of Concern and the reference communities is

- greater than 25%, then the percentage of minority residents in the Community of Concern IS significantly greater than in the reference areas.

For example, a population that is 62% minority would be 55% greater than a reference population that is 40% minority (40% plus 55% of 40%). The percentage of minority residents in this Community of Concern IS significantly greater than in the reference areas.

c. Decision-criterion for Low-Income Factor

The assessment of whether a community is low-income will similarly be based on a comparison of the Community of Concern with one or more reference communities, utilizing the same analytical methodology for each. For EJ purposes, a low-income community is defined as one for which the percentage of household incomes beneath the poverty line specified in Section 4.c., is significantly greater in the Community of Concern than in the reference community, after accounting for the accepted uncertainty in the data. The following guidelines are used for this assessment:

- If the relative difference in the low-income percentages between the Community of Concern and the reference communities is greater than 25%, then the percentage of low-income residents in the Community of Concern IS significantly greater than in the reference areas.

- For example, a population that is 59% low-income would be 31% greater than a reference population that is 45% low-income (40% plus 31% of 45%). The low-income population in this Community of Concern IS significantly greater than in the reference areas.

Because of the complexity in reporting income status, it may be necessary in a limited number of cases to calculate an area's income status in other ways (e.g., unemployment, or labor demographics, and/or level of education), and to look at the results together in a group. A clear consensus would lead to the conclusion that the low-income criteria are satisfied. Anything less than a consensus, or data available for only one or two calculations, would reduce the confidence in any conclusion concerning the low-income.

d. Decision-criteria for Disproportionate Adverse Burden Factor

The assessment of whether a community suffers from a disproportionate and adverse environmental burden will be based on a comparison of the Community of Concern with one or more reference communities, utilizing the same analytical methodology for each. Frequently, a community will have highlighted one or more specific environmental concerns. These should be included in the burden analysis to the extent data and methodologies are available to evaluate such concerns. However, the analysis should also include other potential burdens, covering the various possible exposure routes, as well as other potential contaminants or burden types. An analysis should be performed for each environmental burden that is identified.

In order to assess the relative levels of environmental burden in the Community of Concern and in the reference communities, this Interim Policy specifies procedures for estimating the burden in various scenarios by developing an environmental load profile, which serves as an indicator of burden. It then presents specific cut-offs for determining whether the any individual element in the load profile for a Community of Concern is significantly greater than in the reference

communities.

There may well be situations in which the Community of Concern does not have a specific significantly disproportionate and adverse burden, but may have a number of burdens that, together, represent a significantly disproportionate and adverse environmental burden. For that reason the environmental load profile looks at a number of salient characteristics that relate to the environmental burden on a community. The Interim Policy provides flexibility to the staff conducting the analysis to make reasonable judgements about the appropriate weights of the various elements of the load profile for making the final environmental burden assessment. To the extent practicable, such weights should be decided in advance of the full analysis.

With this in mind, the environmental load profile looks at multiple contributing sources to assess whether a community is defined as suffering from a disproportionate and adverse environmental burden. These elements of burden include (i) some known and quantifiable environmental condition, (ii) the relative magnitude of the burden given the population potentially exposed and (iii) some burden due to proximity to a source of contamination or environmental threat.

Within each of these elements, it is necessary to estimate the specific contribution to burden, which in turn can take many different forms, each with its own measurement problems. Some burdens may be clear and relatively easier to estimate, such as chemical exposures or population proximate to potential sources. Some may be much more subjective, such as perceived odors or indirect effects (like traffic associated with a new source). These perceived or indirect burdens are no less real, they are just much harder to quantify. In addition, some burdens may not exist currently, but may be 'expected,' based on proposed construction, siting, etc.

Since methodologies to quantify actual burden are currently not available, the Interim Policy adopts the concept of developing an environmental load profile for a community. The Executive Order governing EJ, addresses situations where minority or low income communities bear a "significantly disproportionate and adverse burden". To assess whether a community meets the EO requirements EPA needs to characterize the relative burden, not the actual burden, and decide if it is disproportionate and adverse.

The Interim Policy advances the concept of an environmental load profile, which is comprised of different elements that when combined will be indicative of relative environmental burden. The profile would provide a representation of the environmental load in the community, not the actual burden. It would be based on salient characteristics or elements that would serve as indicators of environmental burden and could provide a consistent basis for comparison. The profiles can be compared and the salient characteristics (e.g., indicators of air quality and environmental well being) could be weighed by decision-makers to assist in the assessment of whether the COC is an EJ community.

Although this Interim Policy outlines mechanisms for analyzing environmental data for the purpose of determining particular environmental load characteristics, it does not attempt to

determine appropriate weights for these characteristics (e.g., should air emission analysis be given a greater weight than facility and population density characteristics?). Flexibility should be provided to decision-makers with respect to how to use the information from the load profile.

As previously discussed above, decision-criteria have not been developed by EPA for quantifying burden that is "significantly disproportionate." However, the Region has developed decision-criteria, and cut-offs to quantify whether burden for the individual elements of the environmental load profile are significant, and that when combined provide the Region with a method for comparing whether the burden is also disproportionate.

- For example, if the relative difference between the COC and the reference communities for an environmental load element (e.g., air quality indicator) is greater than 50%, then the component of the burden associated with that indicator IS significant when compared to the reference communities.

- On the other hand, a population that is exposed to 52 ppm for some pollutant would have a 30% increase in burden than a reference population that experiences an airborne concentration of 40 ppm (40 ppm plus 30% of 40 ppm). The relative burden in Community of Concern is NOT significantly disproportionate.

Notwithstanding that one or more elements of the load profile may be significant when the COC is compared to reference community(ies), the assessment of whether this constitutes "significantly disproportionate" would depend on a comparative evaluation of <u>all</u> of the various load profile elements. As discussed above, this comparative evaluation would include the appropriate weighting of the individual load profile elements.

The Executive Order also directs federal agencies to prevent and address significant environmental effects that are <u>both</u> disproportionately high and adverse. This Draft Interim Policy also provides criteria for assessing "adverse." If the burden in the COC is considered by a recognized authority to be safe, then it would not be considered to be adverse, even if it is greater than the burden in the reference community. Therefore, when an acknowledged health/welfare standard exists for the burden of concern (for example, an EPA National Ambient Air Quality Standard), this Draft Interim Policy defines the burden as disproportionate and adverse only if the burden exceeds that standard and exceeds the cut-offs set forth in the Policy. However, this may not be the case when there are several environmental burdens or multiple environmental problems underlying the environmental justice concern.

e. Requirements for Documenting Decisions and Deviations.

The Draft Interim Policy contains requirements, specifying how the EJ assessment decisions are to be documented as provided below. The decision document must state as appropriate:

i. Boundaries of the Community of Concern, and how they were selected.
ii. Boundaries of each reference area and how the reference areas and the boundaries were selected.
iii. For each factor (minority, low income and environmental burden) the quantative analyses that were performed and the results of those analyses.
iv. For each factor, how the Decision-criterion was applied for each community of concern-reference comparison and the result of the comparison.
v. The conclusion of the analysis, incorporating all three factors, or any Special Considerations.

Sub-Part II - Implementation Details

4. Environmental Justice Factors and Data Limitations

a. Introduction

As defined in Section 1, an EJ Area or EJ Community must be:

- minority, and/or low-income, and
- disproportionate and adversely burdened.

This section analyzes each of these factors: providing a practical definition; identifying the sources of data that are available for making the assessment; recommending procedures for the use of the data for EJ analyses in Region 2; and identifying data limitations. As stated in Section 3, however, the three factors (minority, low-income, and environmental burden) in the Community of Concern are only important in comparison to the relative magnitude of the same three factors in the reference areas. Therefore, the selection of the reference communities and the analysis of the three factors for these communities are just as important as the analysis for the Community of Concern. In addition, the same parameters that are used for the Community of Concern must be used for the reference communities as well. The selection of the reference communities is discussed in Section 2.b., Step 2.

In addition, this Section summarizes the issue of cumulative exposure to multiple sources of

pollution, which is of critical importance to communities in which more than one exposure source or exposure pathway is present.

b. Definition of a Minority Community or Population

The Headquarters Office of Environmental Justice has defined the term 'minority' for EJ purposes to include Hispanics, Asian-Americans and Pacific Islanders, African-Americans, and American Indians and Alaskan Natives. For EJ purposes, the term 'minority' does NOT address religion or people who might be distinguished by sex, age, culture, or any type of handicap. The Region 2 Interim Policy therefore defines a 'Minority Community' as a community that has a significantly greater population of individuals in the specified minority groups than does a reference community. Section 3 presented the Decision-criterion for determining whether the Community of Concern is a Minority Community for EJ purposes.

The specification of the Community of Concern and the reference populations is not always obvious, and can have a dramatic effect on the results of the EJ analysis. In fact, in any given situation, there may be a number of ways to divide the population into racial and ethnic groups. Examples might be: separating and enumerating each Census group; combining categories, combining racial and ethnic minorities; including transients, or counting them separately; weighting specific units according to the size of the population, or not. (See Zimmerman, 1994 for a summary of the extensive literature in this area)

c. Definition of a Low-income Community or Population

There are a number of different ways by which the Region might assess whether a given community is a low income community for EJ purposes. Most frequently, The Region will utilize the Department of Health and Human Services (HHS) annual poverty guidelines for a family of four in the 48 contiguous states (a simplification of poverty "thresholds" that are updated each year by the Census Bureau). The Department of Housing and Urban Development (HUD) considers a community to be low-income for its housing benefits programs if its median annual household income falls below HHS's poverty guideline. However, Region 2 utilizes the percentage of families below the poverty guideline, since comparisons of the low-income populations in different communities is the issue here.

In special circumstances, the Region may choose to modify the way in which the income data are used to allow for more meaningful Region 2 analyses. The Region might also choose to utilize a number of surrogate measures that rely on the correlation between 'low-income' and 'poverty.' For example, the Region could use available data on housing costs as a means for getting at the effective relative income levels of communities when the more straightforward approaches (actual income data) do not seem to be appropriate, or are not available. Housing value is usually highly correlated with income, justifying its use as a surrogate. Unemployment statistics might also be used as a surrogate for low-income in certain circumstances.

In order to foster consistency in EJ Community identification, and to facilitate the process, it is recommended that the Region utilize the HUD median family income guideline for a family of

four, unless specific circumstances warrant another approach.

This Interim Policy defines a 'Low-income Community' as a community that has a significantly greater proportion of low-income families than does a reference community. Section 3 presented the Decision-criterion for determining whether the Community of Concern is a Low-income Community for EJ purposes.

d. Definition of a Disproportionate and adverse Burden

Disproportionate and adverse burden (also termed disproportionate and adverse effects or impact, in the context of the Executive Order) is an adverse human health or environmental effect on the Community of Concern out of proportion to the level of the same effect felt in the reference communities, and can be either actual or potential. The environmental burden or impact can be related to ambient conditions, a specific source or sources, cumulative or area-wide sources, and/or uneven application of government authorities. Whatever the environmental burden may be, it must be assessd or estimated in both the Community of Concern and the reference communities in order for the disproportionate and adverse of the burden to be calculated.

Identifying the magnitude of environmental burden in the Community of Concern and in the reference communities is not usually a simple process. The nature, extent, and cause of the burden must all be examined in order for the assessment of the magnitude of burden to be valid. The burden can be manifested in several ways, as health effects or exposures, contaminations, or even proximity to potential exposures or sources. The disproportionality could result from a specific source(s), prevailing ambient conditions, and/or cumulative environmental exposure with a single or multiple agents. The burden can be strictly environmental (such as higher concentrations of some pollutant), or it can be administrative (such as when one area has been or may be treated less favorably than another by EPA or another governmental entity), or a combination of the two.

While it would be preferable that site-specific data be gathered, confirmed, and assessed for EJ area assessments because each case is likely to present its own special circumstances, it is recognized that there are not yet tools available to conduct such an analysis. In addition, although

High quality and consistent data are available for the development of the required low-income and minority demographic profiles of the Community of Concern and the reference communities., There are more limited data available for assessing the environmental burden.

However, it is possible to list the general categories of burden that might be expected:

Human Health Effects (for example, Cancer incidence)
Human Exposures (for example, Ambient air concentrations)
Economic Effects (for example, House value depression)
Regulatory Effects (for example, Delayed cleanups)
Ecological Effects (for example, Destruction of habitats)

Aesthetic Effects (for example, Elimination of scenic view)

This Interim Policy proposes to use consistent methodologies for developing an environmental load profile, to represent burden. The approach is to incorporate methodologies into the load profile where there is a defensible method to provide a quantifiable estimate of an elements contribution to burden and there is a data set available for the entire Region. The Region intends to add categories to the load profile as analytical methods and consistent data sets become available.

Regardless of the type of burden, the available data, calculations, and/or estimations must be developed and analyzed by technical experts who can assess their validity, their associated level of confidence, and their relevance to the situation at hand. In addition, it is important to evaluate the Community of Concern and each reference area in the same way, and with comparable levels of confidence, in order to produce meaningful comparisons.

This Interim Policy defines a'Disproportionate and adversely Burdened Community' as a community that suffers from an existing or potential environmental burden that is significantly greater than the burden suffered by one or more reference community. Section 3 presented the Decision-criterion for determining whether the Community of Concern is a Disproportionate and adversely Burdened Community.

. **e. Availability of Data for Low-Income, Minority, and Burden Assessments**

High quality and consistent data are almost always available for the development of the required low-income and minority demographic profiles of the Community of Concern and the reference communities. There are more limited data available for assessing the environmental burden. In some cases, high quality relevant information may also be available for the environmental burden analysis (e.g., TRI and National Ambient Air Quality information). Further, the types of data that will be needed for the variety of EJ area analyses are extremely varied. To ensure the integrity of the analysis it is essential that a common data set is available for the evaluation. As a result, this IP recommends performing analyses only where such data exists.

i. Low-income and Minority Demographic Data and Data Tools

The decennial census administered by the U.S. Bureau of the Census is the most important source of population, income, and housing data used to describe the income and minority demographic makeup of populations for EJ analysis. The Census Bureau issues a number of different data products that make the decennial census information widely accessible to census data users. The Census Bureau not only issues its data in printed publications, but also publishes the data in electronic format for computer analysis. For example, the Census Bureau publishes its data on computer tapes, compact disc-read only memory (CD-ROM) and diskettes. Users may also access data through the Census Bureau's online service, CENDATA, with the appropriate computer hardware and software.

The Census Bureau computer tapes include:

- Summary Tape Files (STFs), which parallel many printed census reports with significantly more statistical or geographic detail not published elsewhere;

- microdata sample files, which include records for unidentified individuals, households, and housing units. Geographic information is limited and the files are published in a format that protects the confidentiality of responses; and

- TIGER/line files, assign geographic codes to addresses for computer mapping.

Demographic and economic information for States, Counties, and places is contained on Census Bureau diskettes. A list of census data available on CD-ROM is provided in **Appendix A**. Of the CD-ROM data files, the most widely used for EJ analysis by EPA are STF 1A and STF 3A, both of which are available down to the block group. STF 1A is 100-percent count data and contains population data, including race, age, income, and housing information. STF 3A is sample data, but is more detailed, including income and education statistics.

The Region 2 Geographic Information System (GIS) contains 1990 Census STF 3A, TIGER/Line data, facility locations from EPA's facility databases, hydrology, soils, land use and other geographic information. For a complete list of data layers available on the Region 2 GIS, please see **Appendix B**. For information about the application of GIS to EJ, please see **Appendix A.**

There will be some circumstances in which the data available in the standard data bases are not sufficient or appropriate for a specific EJ analysis. For example, certain areas in the Region may have large numbers of undocumented or transient residents who are not recorded on any official census data bases. When it is clear that a specific community may have demographics that are significantly different from the official census figures, the Region will need to take more aggressive measures to develop more meaningful data. For example, the Region may wish to conduct or arrange for an informal survey of the area.

ii. Environmental Burden Data (including ecological & health burden)

As stated above, environmental burden can take many forms. The most relevant forms for EJ analysis - direct human health effects - are usually the hardest to document. They will be available only if a specific study has been conducted in the area, and then may not be available for the reference areas. State and local health departments compile health data that are available to researchers. For example, the New York State Department of Health (NYSDOH) collects cancer, infectious and heart disease statistics for New York State, as well as vital statistics, including births, deaths, and spontaneous fetal death. The Centers for Disease Control and Prevention (CDC) publishes similar data for the country, including health

risks (e.g., behavioral risk factors, environment and health). In addition, hospital and emergency room data are available for certain conditions (like asthma) in certain areas. Because of the lack of a consistent data set for this information, the Region anticipates using this type of data only in special circumstances.

Information on exposure may be found, at least generally or indirectly, in some commonly available data bases. For example, there is a large amount of facility (source) information available for EJ analysis in EPA's mainframe databases, such as the Toxic Release Inventory System (TRIS), Resource Conservation and Recovery Information System (RCRIS), Permit Compliance System (PCS), Comprehensive Environmental Response Compensation and Liability Information System (CERCLIS), Federal Reporting Data System (FRDS), and the Aerometric Information Retrieval System (AIRS) Facility Subsystem (AFS). Ambient Monitoring data may be obtained from AIRS and the Storage and Retrieval of Water-Related Data System (STORET). However, these data bases only address certain categories of facilities and pollutants, are not complete, and are not all of consistent quality.

There will be instances in which the health data of interest will not be available at all, or only for the Community of Concern but not for the reference community. The quality of the environmental burden data is highly variable. Some ambient air and water quality measurements are very reliable, while some may be sketchy or non-existent. Most information about specific health effects, exposures, loadings, and government interventions will not be found on any standard data base, and will need to be developed as needed, frequently utilizing models that bring new layers of uncertainty to the analysis. As stated above, the available data, calculations, and/or estimations must be researched, developed and analyzed by technical experts who can assess their validity, their associated level of confidence, and their relevance to the situation at hand.

f. Cumulative Exposures

Concerns regarding exposures to multiple contaminants through multiple contamination pathways are central to many EJ assessments. However, characterization of cumulative exposure from multiple sources is problematic for several reasons. Most significantly, EPA tends to regulate emissions in a site-specific and statute-specific manner, and releases from non-point sources or uncontrolled sites are exceedingly difficult to quantify. While comprehensive geographically-based data on chemical releases and potential exposures are difficult or impossible to assemble, some valuable information may be generated from currently available data sources.

The Toxic Release Inventory (TRI) represents one of the most important readily available estimates of emissions of toxic compounds from large industrial facilities. The TRI lists annual releases of each of 340 toxic chemicals from large quantity generators to air, water, or land. While the TRI is not comprehensive (many facilities are not covered, such as small-quantity generators and uncontrolled hazardous waste sites, and many chemicals are not reported), the TRI does provide valuable information for use in a screening survey for areas of potentially elevated exposure, particularly through Geographical Information Systems (GIS). Numerous GIS

applications (environmental targeting systems) have been developed for the analysis of TRI data. A compendium of these systems is presented in U.S. EPA 1994.

The majority of these GIS/TRI systems are appropriate for screening-level studies, as follows. First, the quantity and toxicity of the reported releases are frequently integrated to generate an index of the potential hazard associated with the release. The resulting information is mapped to specific geographical locations, sometimes along with factors related to the potential dispersion of the contaminants released. This information is then generally combined with various demographic factors to generate maps that can enhance understanding of where certain human populations (such as minorities) may be experiencing potentially significant exposures.

Perhaps the greatest challenge in the development and use of GIS/TRI applications is to maximize the information available from the data without generating analyses that lack meaning through over-interpretation of the data. Still, the Region believes that a TRI, GIS based tool could be developed to provide for a comparative analysis that would allow for an assessment of disproportionate and adverse risk.

5. Geographic/Boundary Definitions & Consideration

a. Introduction

Any assessment of whether an existing or potential environmental burden constitutes an Environmental Justice burden depends first on the clear definition of the areas under consideration. The analysis of economic, population and/or burden factors cannot even be attempted until the geographical boundaries of the Community of Concern and the reference areas have been clearly considered.

There are a number of different ways in which the geographic boundaries of the Community of Concern and the reference areas might be identified. For example, one logical approach would be to define boundaries according to established political boundaries - city, county or town limits. Another would be to utilize obvious physical boundaries, such as rivers, main roads, or railroad tracks. A third approach would be to utilize the U.S. Census boundaries, such as those for census blocks or block groups. This procedure is the most convenient for GIS mapping purposes, and also takes advantage of the fact that the Census tries to match its blocks and block groups to the realities of the population demographics. As such, an approach based on Census compilations would normally be the most relevant and useful for the Region. However, it is not always clear how the Census areas should be grouped - or whether a special case exists that mitigates against using the Census compilations at all. An important approach for learning about the special boundary considerations in a given Community of Concern is to utilize community input. Not only can this produce valuable information, but it can also provide a vital link to public participation in the Region 2 EJ process.

The remainder of this Section addresses the use of community input for better defining boundaries for both the Community of Concern and for the reference areas, followed by a more detailed explanation of the types and utility of various Geographic Units of Analysis for defining exposure areas. For most purposes, the same techniques are used for defining the Community of Concern and reference boundaries. When they are different, the text will so state.

b. Types of Areas

i. Community Based Definition for Community of Concern and Reference Areas

Community input is a key factor in the identification of EJ Community of Concern and reference areas. Residents generally can identify the boundaries of their community. They often can identify the wealthier and poorer areas, and generally know whether the community in the shadow of the factory is a minority community. EJ boundaries should, in large part, follow the boundaries of communities, that is, more or less intact groupings of people in contiguous geographic areas for which the boundaries are acknowledged by the community members themselves. For example, the boundary lines surrounding Bushwick in Brooklyn or Ironbound in Newark may not carry the political weight of the New York City or Newark city limits, but, for EJ and other social purposes, they may be much more important.

Two excellent sources of information on community boundaries are the community's governmental leadership (municipal government, for example), and the communities themselves (community action groups, etc.). This latter approach might be particularly valuable when a minority/poor community is at odds with the larger political entity. While it would not be possible for the Region to consult with each and every community and community group within the Region, it will be possible for the Region to be responsive to, and even to seek out specific communities when there might be a dispute concerning the appropriate boundaries for EJ purposes.

There are four reasons for the Region to take advantage of this acknowledged set of community boundaries for its own EJ analyses: first, it makes some decision-making easier; second, it is likely to add a great deal of practical reality to the process; third, it will enhance the Agency's level of cooperation with the communities; and fourth, it will promote equity in the process of Agency decision-making (which is also written into many environmental statutes).

It may also be very beneficial to gather this information through discussion with state and local authorities, many of whom already have established community boundary lines based on practical day-to-day interactions with the communities. They may well have maps, even GIS-compatible data files, that the Region could use. In response to specific situations, and for the Region as a whole as time allows, the Region will begin the process of consulting with state and local officials to obtain their recommendations for drawing community boundaries. This process will also further the vital EJ principle of building partnerships with state and local government agencies, who are certainly stakeholders along with the communities.

While the individual circumstances may vary, there are likely to be some consistent techniques for enlisting community assistance in determining appropriate boundaries for EJ areas. For example, it is likely that the Region will meet with the leaders of local community groups to discuss this issue, probably in the context of larger EJ issues related to the project or problem at hand. In more controversial cases, or cases in which there are various community groups, it may be beneficial to discuss this issue in a small 'town meeting' or other public forum.

An additional potential benefit from discussing community boundaries with local officials and/or community members is learning information about special populations that might not be included in the available population reports. These reports usually focus on residents (or, less commonly, people in other fixed facilities, such as workplaces or institutions). They usually do not address transients, such as recreationists, tourists, shoppers, students, or other visitors and occasional workers. The presence of significant numbers of such individuals might well require special surveys to ensure their inclusion in the EJ analysis.

ii. Boundary of the Reference Areas

EJ is intended to address situations in which specific populations (minority and/or low-income) have suffered (or are projected to suffer) greater environmental burdens than other (non-minority and/or non-low-income) communities. The analysis of potential EJ areas requires the identification of reference areas in order to provide a context for interpretation of data from the Community of Concern.

There are two general types of reference communities. In the first case, the reference community is sufficiently proximate to the Community of Concern to render the analysis relevant. For example, the reference area may be adjacent to, or may contain the Community of Concern (and the area affected by the burden). The second type of reference community is needed when the environmental burden in question is the result of a site-specific government decision (as when EPA may have cleaned up one community before another, or when a permit is granted to a facility in one community but not in another). The reference communities may then be those areas that were not selected as sites for the environmentally burdening facility/activity, or were selected for allegedly preferential clean-up.

For practical purposes, potential reference communities must be proposed at the start of the analysis in order that all of the minority, income, and burden analyses can be performed objectively. This proposal of reference communities will be based on readily available information, screening analyses, and prior knowledge of the burden, the service area of any relevant source or potential source, and the area proximate to the Community of Concern. Once the detailed analyses are underway, one or more of the proposed areas may become clear as the reference area of choice. However, it will usually be a good idea to continue the analysis for two or more potential reference areas, as disputes often arise concerning the selection of the reference area, often related to interpretations of the demographic data, or the relative mix of environmental burdens.

If possible, it is important to choose communities that are sufficiently close (proximate) and/or comparable to the Community of Concern that it would be reasonable to assume the presence of similar circumstances if EJ were not a factor. For example, if the Community of Concern is an urban poor community with many bus terminals, a rural community across the river in another state would not likely satisfy any logical claim to being a legitimate reference community.

Before selecting a reference community(ies), selection criteria should be developed. That is the particular community features that are associated with the EJ issue should be identified. Those features should be used as selection criteria for identifying appropriate reference communities. For example, in the case of siting a facility, issues such as proximity to raw materials, transportation routes, etc. may be relevant requirements for the proposed location of a facility. The COC and reference communities should have those similar features.

A reference area should be well outside the area affected by the burden that prompted the EJ concern, unless the burdened area is very small or very large compared with the size of the reference area. When the Community of Concern is contained within a larger geographic area that is used as the reference area (such as county or state), the reference area should be large enough to contain a number of discrete areas, particularly when more than one source is located within the reference area. On the other hand, when the environmental burden covers much more than the Community of Concern, the reference areas may well be selected from within the burdened area.

When a general screening is being conducted, such as to find "all of the potential EJ areas in New Jersey," the entire area in question serves as its own reference area, and all specific areas are compared against all others, usually in terms of the two demographic factors only. Those areas with the highest minority and/or low-income percentages are then identified as having the highest potential to be EJ areas. It is important, though, that such a screening exercise be considered as such, and that the results not be treated as a final analysis. As long as the results are used only for general purposes, or to target more specific and detailed analyses, screening exercises can be extremely valuable.

The ways in which the boundaries of reference areas are defined and the use of data units are similar to those that will be outlined under Communities of Concern. One other method for delineating a reference area is often used - a ring of data units around the Community of Concern, using data units that are similar in type to the ones used to delineate the Community of Concern. The chosen reference areas should be of varying size, from areas similar to the Community of Concern to those that are very large, such as a state.

iii. Source Boundary as Input to the Boundary Analysis

The first step in an EJ analysis involving known sources of potential pollution is to define the

location of the sources. Sources of adverse burdens, whether facilities, sites or area sources, are typically defined as points delineated by a latitude and longitude, preferably measured by GPS, or located via areal photographs. A useful alternative, however, is not to narrow the location to a single point but rather to define it as an area. To save time, and when census data are used for demographics, that area is usually taken as the same Census data unit used for the demographic information (e.g., Tract, Block Group). The use of the Census Tract has been very common for this purpose, especially when detailed latitude and longitude information for the source boundaries is not available and only the Tract assignment is available. The advantage of using the Tract is that there is a very rich socioeconomic data base for such data units. Moreover, in dense urban areas, the Census Tract is often smaller in area than a circle circumscribed by a 1-mile radius, (another common alternative).

A disadvantage of using the Tract is that without knowing where the site is actually located within the Tract, one could miss population characteristics that are very near the site, but are located in a different (adjacent) Tract. This occurs when the site is not located in the center of the Tract. Second, if the area of the site is very large and even larger than the area of the Tract, then the Tract would not be a good representation of a Community of Concern. In such cases, the Community of Concern would be adjacent Tracts. Third, in areas that are not densely populated, the Tract is usually much too large to serve as a meaningful Community of Concern.

iv. Boundary of the Burden as Input to the Boundary Analysis

Ideally, modeling for environmental fate, transport, and exposure should be conducted to define the boundaries of the impacted environmental media (air, water, land, etc.). Since it is usually difficult if not impossible to define these areas, and detailed risk assessments are usually required for such purposes, surrogates or approximations of the burden boundaries are used.

Several alternatives have been used as surrogates for delineating a potential area under environmental burden. These are: a Census data unit or combination of units, usually the Census Tract around the facility (or the Data Unit within which the facility is located); or some circular or approximately circular area around the facility (either obtained through Geographic Information Systems (GIS) or a “centroid pull” technique, which includes all of those data units whose centroids fall within the specified radius).

Circular areas are commonly used to represent environmental burden boundaries for EJ analyses at Superfund and RCRA sites, particularly when a number of sites are to be studied at one time. Alternative radii include 0.5, 1.0, 2.0, 3.0 and 4.0 miles. A 4.0 mile radius has often been used in exposure modeling. Circular areas have the advantage of introducing small and consistently defined distances from the various sites. The characteristics of areas using alternative radii can be compared easily with one another. A potential disadvantage is that the circumscribed circle may not be sufficiently large to characterize the burdened area, or the burden may not be equal in all directions from the source. It is generally best to choose a

number of different criteria for defining the burdened areas and comparing the results.

There are two important concepts that must be remembered when determining the Community of Concern and the reference areas. First, the reference areas may actually lie within an environmentally burdened area, even the same one as the Community of Concern. Second, the Community of Concern may be determined using the same surrogate techniques described here for determining the boundaries of the environmental burden, with appropriate input from the community.

c. Data Units

i. Census based

Except where survey data exist for individuals, population data are usually only available for discrete units in order to safeguard the confidentiality of the data. As stated above, the most common source of information is the U.S. Bureau of the Census, Census of Population and Housing, which is published every ten years with population updates every few years. When Census data are used, the choice of data units (roughly from smallest to largest) is: Blocks, Block Groups, Tracts, zip codes, various metropolitan area designations, Counties, and States. Additionally, special areas exist such as tribal areas.

The criteria for selecting a given Census unit are generally driven by the selection of the population subgroups and by the availability of desired data for the particular data unit. The size of the data unit should be as small as possible in order to increase the accuracy of describing population characteristics within a given boundary. Although Census data are clearly aggregated at larger geographic units, the larger the unit the more difficult an EJ analysis becomes. Since Block data is the finest, it should be used where possible. Although the choice of parameters available at the Block level on CD ROM is limited (data on income are not available in this way, but data on house value are), a more detailed set of parameters is available on tape. For most EJ analyses, more detailed socioeconomic characteristics may not be necessary, and may not be worth compromising the spatial resolution.

The geographic differentiation of environmental or health information (when available), however, should be kept in mind when selecting the geographic unit for population characteristics. That is, it may not be useful to have data units for population characteristics which are smaller than the geographic units for environmental or health data, and vice versa.

Examples: Block data have been used for Superfund sites (Zimmerman (1994, 1996) for EPA Region 2 and by other EPA Regions and Headquarters). Block groups are more commonly used and were used in conjunction with TRI data by Perlin, et al. (1995) and for the majority of work by EPA regions. Tracts are also commonly used, and have been used in EJ analyses of RCRA, NPL, TRI and general environmental conditions. Zip Codes were used as data units in some of the earliest EJ studies (United Church of Christ, 1987; Center for Policy Alternatives, 1994). Municipalities (usually defined by the Census as Places, Minor Civil

Divisions, Cities, Metropolitan Statistical Areas, etc.) were used by Zimmerman (1993) and Anderton, et al. (1994, 1995). Counties have been used rarely, although one study of Superfund sites was conducted by Hird (1993). A more detailed set of examples may be found in Appendix C.

ii. Other

Direct surveys of population within the boundaries of interest are perhaps the best way to obtain population information. While this is the a very costly and time consuming approach, it may be necessary for collecting information on groups not well captured by the Census. It may also be possible to collect sufficient information on various transient groups like students or workers through their schools, employers, unions, etc. In any case, when significant numbers of such individuals are believed to in the Community of Concern and potentially affected by the environmental burden, or in the reference areas, it is important that they be considered fairly.

6. Evaluation Process

a. Introduction

Sections 4 and 5 have described how to define minority, low-income, and environmentally burdened communities, and how to define the Community of Concern and the reference areas. However, for the demographic factors (minority and low-income), it is also necessary to define how the data are actually extracted from the available sources, and how they are prepared for the comparisons of the Community of Concern and the reference areas. The rest of this Section will describe methods for extracting and transforming demographic data, and appropriate statistical procedures.

b. Extracting Population Data

The best and most direct means for obtaining total population and population subgroup counts would be by conducting direct surveys or head counts of the residents and transient populations within the boundaries of the area of interest. Since this kind of data rarely exists, and is difficult and expensive to obtain, EJ analyses most often utilize population data collected by the U.S. Bureau of the Census as a surrogate. Population characteristics from the Population and Housing Census are available in a variety of alternative geographic units: blocks, block groups, tracts, metropolitan areas, counties, states, etc., from which the characteristics for the Community of Concern and reference areas have to be "extracted."

The first step is to decide which data unit to use. The decision is based upon the availability of the data for the areas under consideration, and the size and shape of the Community of Concern and the reference areas. Since the discrete Census geographic data units may be smaller than or not coincide with the areas of interest, the data in the data units must usually be aggregated (grouped) in order to characterize the population characteristics for the areas of interest.

c. Available Demographic Data Extraction Techniques

The selection of techniques for data extraction and aggregation will normally depend on whether the aggregation of the data units is needed to conform to specific, pre-defined boundaries of the Community of Concern or the reference areas, or whether an approximation is appropriate, as when circular areas are to be used..

- Conformation to a predetermined boundary: Geographic Information Systems (GIS) techniques should be used. GIS techniques allow the aggregation of population by intersecting the appropriate contiguous census units to fit a predefined geographic area. This inevitably requires some assumption about the population in the units being intersected, most commonly that the population is homogeneously distributed across the data units. The advantages of GIS are that the aggregated data units can be made to conform to a specified boundary, and the data can be mapped easily.

- Conformance to pre-determined boundary not needed: Census Centroid Pull (CCP) technique should be used. This technique aggregates the population within the data units whose centroids fall within a pre-determined distance (radius) from a point. The outer boundary of the geographic area consists of the combined outer boundaries of the included data units, rather than the smooth edge of a circle. Although GIS may be used for this approach, CCP is less expensive and quicker when data units need not be intersected.

The details of the GIS and Centroid Pull techniques are described in the Appendix A. GIS is the most commonly used data extraction technique in studies such as these, and is currently being used by EPA Regions, typically with block group or tract data. The Census Centroid technique was used by Zimmerman for NPL sites (1994), and for non-NPL sites (1995) for Region 2 Superfund site analyses.

d. Data Transformations (including aggregation)

i. Data Transformation for Consistent Data Sets for the Community of Concern and the Reference Areas

Population and environmental characteristics are described and quantified in a number of different ways. In order to compare characteristics of the Community of Concern and the reference areas, the data have to be put into a common and usable format. Some of the data may have to be transformed to make the data sets consistent. When such data changes or transformations are undertaken, it is critical to avoid biases produced by data alterations. The following discussion of data transformations pertains to Census data utilized for the population characteristics. Data transformations are also commonly needed for the environmental data. However, because the data will often take widely differing forms and formats, it is not practical or even possible here to address the transformation process effectively. Suffice it to say that the transformations and aggregations of environmental data to be utilized in EJ assessments must be performed by qualified professionals, and must be thoroughly documented and described. The same warnings concerning the inadvertent insertion of bias into the analysis will, of course, apply.

In any evaluation of population data, it is assumed that the basic geographic unit for which data are obtained is a Census unit, such as a block, block group or tract. Whichever Census data unit is used, both the Community of Concern and the reference areas are usually comprised of numerous such data units, requiring that the data from a number of data units must be aggregated (combined) to produce usable information for each.

There are a number of different ways to utilize population data from the Census:

- *Raw data* may be used, usually the number of people in each specified category living within each of the Census data units in the Community of Concern or the reference areas (such as the number of low-income individuals in each block). This produces the absolute difference in the numbers of persons in the Community of Concern compared with each

reference area. The difference, however, may only be due to the absolute difference in the total numbers of people in the two areas (or the population densities) rather than to any difference in the proportion of minorities.

- The raw data may be converted into a relative scale such as a *percentage or a proportion* by dividing the number in each category in each area by the total number of individuals in the area. This comparison filters out differences in population size or density, and is the common form used for most population characteristics.

- Either raw data or percentages may be used to construct summary statistics - usually as *averages* - to characterize the typical attributes for the Community of Concern and the reference areas. Examples of commonly used summary statistics are means, medians, and modes. These statistics are computed by aggregating the raw data or percentages for each of the data units, and may be either weighted or un-weighted by the population in each of the data units (see the discussion of weighting factors in the following sub-section).

- Additional summary statistics, such as *distributions*, may be constructed to clarify the nature of the population characteristics across the data units within the Community of Concern or the reference areas, including the identification of extreme values or outliers. These are typically expressed as ranges, quintiles, quartiles, or other subdivisions of the data.

- Most forms of the data described above may be transformed further into *indexes*. Values for the data may be assigned to intervals (such as, for percentages, intervals of 10%, ranging from 0-100%). Each interval may, in turn, be assigned a number (for example, ranging from 1 to 5) or some other counting mechanism (such as quartile) in order of increasing magnitude of some population characteristic. These indexes are useful because they can help to make complicated or confusing statistical comparisons more manageable without losing much informational value.

ii. Some Considerations in Transforming Data

Consider Weighting Factors. There are two ways to construct averages for the aggregated data on the population characteristics in the Community of Concern and the reference areas, and they give different results. It is important to decide up-front which method will be employed. The first, and most common, method is called weighting. In this technique, all of the people in the group of concern - low-income or minority - are summed across all of the individual data units within the area, and that number is divided by the sum of all of the people within the same total area. The result is the fraction of low-income or minority people within the total population. It is called the weighting method because it weights each data unit according to that data unit's total population. In the second method, non-weighted, the fraction of low-income or minority people is taken for each data unit, and those fractions are then averaged. This method is called non-weighted because it gives each data unit exactly the same weight. The weighting method should be used here unless there is a compelling reason to do otherwise.

Avoid Bias in Indexes. Some transformations, particularly indexes, can produce considerable distortions if not performed carefully. Indexes can be particularly valuable in screening-level presentations upon which specific decisions will not be based. However, the use of indexes should be avoided in earlier steps of a detailed site-specific EJ assessment. When indexes are to be used, there are a number of ways to minimize the introduction of bias. First, construct indexes only after aggregating the raw data, not before. Second, select break points (or intervals) carefully, so as not to grossly distort the frequency distribution. Also, indexes cannot be used for statistical correlations and regressions, or when simple statistics such as means and medians are to be constructed from the data. Finally, perform a sensitivity analysis to determine how much the results change with different formulas for constructing the indexes (such as different break points).

Include Estimates of Uncertainty. Regardless of which summary statistic or aggregation technique is used, it is important to include estimates of uncertainty, such as standard deviation or other measures of data quality. A major consequence of the use of uncertainty estimates is that they produce ranges instead of specific values. This may make decision-making more difficult, but it gives the decision makers an opportunity to better understand the level of confidence they should have in their decisions. In practical terms, the inclusion of uncertainty estimates in the demographic and/or environmental burden calculations will require that the differences between the Community of Concern and the reference areas be even greater in order for the Community of Concern to be judged an EJ community.

e. Statistical Analyses

Once values are obtained for a given characteristic for the Community of Concern and the reference areas, conventional statistical techniques are used for determining whether or not the differences are statistically significant. Statistical techniques, such as correlations, may be used to determine the degree to which population differences are associated with differences in environmental conditions. In addition, statistical tests should be performed after the Decision-criteria have been applied to confirm that appropriate reference areas were used.

A difference-of-means test is appropriate for comparing averages (means, medians) for characteristics of the Community of Concern and the reference areas when these averages have been obtained by aggregating data for numerous data units. This test assesss whether or not the average of a given characteristic of the Community of Concern is statistically significantly different from the average for that same characteristic in the reference area. It is performed for each characteristic individually. A level of significance has to be specified ahead of time for determining the significance of the difference (e.g., probability that they are different is .01, .05, etc.).

When the Community of Concern and/or the reference area comprises a single Census data unit, then a simple comparison between the characteristics of the two areas may be made.

Before the Decision-criterion may be applied for each EJ factor, the appropriate statistical test must be conducted to assess if the observed difference between the Community of Concern and the reference areas is real. If a statistically significant difference is observed, then apply the Decision-criterion to that value. However, if a statistically significant difference is not observed, then the apparent difference between the Community of Concern and the reference areas for that factor is considered to be zero for the purpose of applying the Decision-criterion for that EJ factor.

Finally, sensitivity analyses or more qualitative assessments should be performed after the Decision-criteria have been applied, to ascertain the difference in results associated with the different reference area selections. Also, a comparison should be made of the robustness of the results of each.

GLOSSARY

Adverse Environmental Burden	A harmful environmental burden. When there is an acknowledged health or welfare standard for the burden in question, the burden is adverse only when it exceeds that standard. When there is no standard, the decision is more subjective.
Aesthetic Effect	An environmental effect based on desirability in appearance, taste or odor, but not associated with adverse impacts on health or welfare.
AFS	USEPA AIRS Facility Subsystem
Agent	The substance that is responsible for an impact on health or welfare.
Aggregation	Grouping of several discrete sets of data, such as combining populations across several sub-areas**.**
AIRS	USEPA Aerometric Information Retrieval System, the database containing the Agency's air-related data.
Ambient Condition	The meteorologic or atmospheric state in a specific location.
American Indian	All indigenous populations within the Region, regardless of their affiliation with a federally-recognized Tribe.
Bias	A systematic or subjective distortion of statistics as a result of the sampling procedure or interpretation.
Block	Census blocks are small areas bounded on all sides by visible features such as streets, roads, streams, and railroad tracks, and by invisible boundaries such as city, town, township, and county limits, property lines, and short, imaginary extensions of streets and roads.
Block Group	A unit for census data reporting formed by a cluster of census blocks. Census block groups generally contain between 250 and 550 housing units.

Boundary	A limiting or bounding line between two or more geographical areas.
CD-ROM	Compact Disc - Read Only Memory
CENDATA	U.S. Census Bureau online database
Census	An official enumeration of the population, with details as to race, age, gender, income, etc.
Centroid Pull	A procedure for grouping census data; all data units (such as block groups) with centroids (geometric centers) that fall within a defined radius from a central point are included in the analysis.
CERCLIS	Comprehensive Environmental Response, Compensation and Liability Information System, a database containing information on Superfund sites.
Chi-square	A statistical test for determining the mathematical fit of a frequency curve to an observed frequency distribution.
Community Input	Information provided by representatives of an affected community on neighborhood boundaries, health concerns, etc.
Community of Concern	A community that is the subject of an Environmental Justice analysis.
Confidence Interval	A range of values for a specific parameter that is believed, with a preassigned degree of confidence, to include the true value.
Contiguous	Bordering or adjoining (as in a neighboring community).
Criteria	Established standards for environmental contaminants.
Cumulative Exposure	Total exposure to environmental contaminants, including exposures originating from multiple sources.
Data Layer	An input parameter for a geographic information system (GIS) analysis covering the area of concern. Information about topography, roads, population, industry, or pollution

might all be separate data layers that may be combined in a GIS analysis to present a useful picture of the area on a single map.

Data Transformation The process of converting data into a format that is more readily used in an analysis. May include conversion into percentages or intervals, or use of the data to generate summary statistics, such as the mean or median.

Decennial Census The 10-year official counting of the U.S. population, with details as to gender, age, income, etc.

Decision-Criterion An established test used for determining whether a Community of Concern meets a specific EJ factor (such as low income, minority etc.).

Demographic The statistical data of a population, especially regarding race, ethnicity, gender, income, etc.

Difference of Means Test A statistical test designed to determine whether the average values of two data sets are significantly different.

Dispersion Spreading widely; scattering or sending off in a variety of directions and distances.

Disproportionate Burden Disproportionate Environmental Burden

Disproportionate and Adverse Environmental Burden
The adverse human health or environmental effect on a particular community or segment of the population (the Community of Concern) that is out of proportion to the level of the same effect felt in reference communities. The burden can be related to a specific source or sources, resulting from cumulative or area-wide sources, and/or resulting from uneven application of government authorities.

Effective Relative Income A ratio representing the comparative annual wages of two or more populations.

EJ Environmental Justice

EJ Area A minority and/or low income area suffering a

	disproportionate and adverse environmental burden as a result of the unfair or unequal development, implementation, or enforcement of environmental laws, regulations or policies (the same as an EJ Community or EJ Population).
EJ Community	A minority and/or low income community suffering a disproportionate and adverse environmental burden as a result of the unfair or unequal development, implementation, or enforcement of environmental laws, regulations or policies (the same as an EJ Area or EJ Population).
EJ Assessment	Use of this Interim Policy to evaluate whether a specific community in EPA Region 2 is an EJ Area.
EJ Population	A minority and/or low income community suffering a disproportionate and adverse environmental burden as a result of the unfair or unequal development, implementation, or enforcement of environmental laws, regulations or policies (the same as an EJ Area or EJ Community).
EML	USEPA Exposure Models Library
Environmental Burden	The adverse human health or environmental effect on a particular community or segment of the population related to a specific source or sources, resulting from cumulative or area-wide sources, and/or resulting from uneven application of government authorities.
Environmental Impact	Environmental Burden or effect.
Environmental Justice	The fair treatment and meaningful involvement of all people regardless of race, color, national origin, culture, or income with respect to the development, implementation, enforcement and compliance of environmental laws, regulations, and policies. Fair treatment means that no groups of people, including racial, ethnic, or socioeconomic groups, should bear a disproportionate share of negative environmental consequences resulting from industrial, municipal, and commercial operations or the execution of federal, state, local and tribal programs and policies.
Environmental Load Profile	A representation of the environmental load in a community, which is based on salient characteristics or elements that serve as indicators of

environmental burden and provide a consistent basis for comparison.

EO	Executive Order 12898 "Federal Actions To Address Environmental Justice In Minority Populations and Low-Income Populations," issued by President William J. Clinton on February 11, 1994.
Ethnic Group	A group of people of the same race or nationality who share a common and distinctive culture.
Eutrophication	An increase of mineral and organic nutrients in a body of water, resulting in a reduction of dissolved oxygen and creation of an environment that favors plant over animal life.
Exposure	Subject to the action or influence of environmental contaminants through ingestion, inhalation, or skin contact.
Extraction	To collect information from a database, such as the U.S. Census.
Facility	A factory, plant, industry, utility, or commercial establishment that is a potential source of environmental contamination or degradation.
Federal Register	Publication of the U.S. Government listing government announcements, rules and regulations.
FRDS	Federal Reporting Data System
Geographic Information System	An organized computer system designed to efficiently capture, analyze and display forms of geographically referenced information. Commonly, GIS is used to combine various data layers (for example, population demographics and environmental burden) to produce maps that display the layers together, allowing for convenient visual analysis.
GIS	Geographic Information System
Government Authority	A local, state, or federal governing body that has the authority to act in environmental matters.

Government Intervention	Judgements, actions or commands taken by a local, state, or federal governing body to address a specific environmental issue.
HHS	U.S. Department of Health and Human Services
Hispanic	Persons who classify themselves as Mexican, Puerto Rican, or Cuban, as well as those who indicate that they are of other Spanish/Hispanic origin.
Housing Value	A data category in the U.S. Census that represents the attributed worth of the homes in a designated area; often used as a surrogate for income in demographic analyses.
HUD	U.S. Department of Housing and Urban Development
Hydrogeology	The science that deals with the occurrence, circulation, distribution, and properties of the water of the earth and the earth's atmosphere.
IMES	Integrated Model Evaluation System
Indexes	Numbers or formulae expressing some property, ratio, etc**.** or used to characterize a set of data in a simplified manner.
Infectious	Communicable by infection, as from one person to another.
IPS	Interim Policy Subgroup of the Region 2 Environmental Justice Workgroup
Large Quantity Generator	A facility that produces greater than a threshold quantity of a substance annually (i.e., greater than 1000 pounds of a specific chemical substance).
Level of Confidence	A measure of the degree of certainty in a statistical conclusion.
Low-Income	Having an annual income that is less than a pre-assigned cut-off. The Interim Policy utilizes the U.S. Department of Health and Human Services poverty guideline as the cut-off.
Low-Income Community	A community that has a significantly greater population of low-income families than does a reference community.

Mainframe	The device within a large computer that contains the central control and arithmetic units.
Mean	The average value of a group of values.
Median	The middle number in a given sequence of numbers.
Microdata Sample File	A database from the U.S. Department of Census that includes records for unidentified individuals, households and housing units.
Minority	An individual or group of individuals that are Hispanic, Asian-American and Pacific Islander, African-American, American-Indian or Alaskan Native. (For EJ purposes, the term 'minority' does not address religion or national origin. It also does not include people who might be distinguished by sex, age or any type of handicap).
Minority Community	A community that has a significantly greater population of minority individuals than does a reference community.
Mode	The observation in a distribution that occurs with the greatest frequency.
Multi-variate Analysis	Statistical analysis of the probability distributions of two or more discrete random variables.
NEPA	National Environmental Policy Act
NPL	National Priorities List
NYSDOH	New York State Department of Health
OEJ	USEPA Office of Environmental Justice
PCS	USEPA Permit Compliance System, a database of water dischargers.
Percentage	A rate or proportion per hundred.
Percent Poverty	The percentage of household incomes in a community that fall beneath the poverty line specified by the Department of Health and Human Services.

Political Boundary	The line dividing two areas with separate governing bodies, i.e., cities, counties or states.
Population Density	The number of people contained with in a defined unit area, i.e., persons per square mile.
Proportion	The comparative relation between things or magnitudes; ratio.
Proximate	Adjacent, or very near to.
Quartile	In a statistical frequency distribution, one of the values of a variable that divides the distribution into four groups having equal frequencies.
Quintile	In a statistical frequency distribution, one of the values of a variable that divides the distribution into five groups having equal frequencies.
Racial Group	A group of persons related by common descent, blood, or heredity.
Range	The difference between the smallest and largest values in a statistical distribution.
.	
Raw Data	Data that have not been transformed or manipulated in any way.
RCRA	Resource Conservation and Recovery Act
RCRIS	USEPA Resource Conservation and Recovery Information System
Reference Community	A community or area selected for comparison with a previously identified community (the Community of Concern) for determining whether the Community of Concern is out of the norm with respect to demographics or environmental burden.
Regression Analysis	A statistical method used to estimate the value of a variable from a knowledge of the values of one or more other variables, and the measurement of the errors involved in this estimation procedure.
Regulatory Effect	A potential form of environmental injustice characterized by

	bias in the administration of government programs.
Robustness	In referring to a statistical test or measure, describes the test's lack of sensitivity to non-normality in the data being analyzed.
Screening Analysis	An initial geographical and demographical analysis for identifying potential Environmental Justice areas or sites that may pose Environmental Justice concerns.
Sensitivity Analysis	A statistical procedure conducted to identify the factors in an analysis that have the greatest bearing on the outcome.
Site-Specific Analysis	An analysis intended to assess whether a specific identified site or area (the Community of Concern) is an Environmental Justice area or poses Environmental Justice concerns.
Skewness	Asymmetry in a frequency distribution.
Source	The site or facility from which a particular suite of environmental contaminants originate (i.e., factory, incinerator, etc.).
Standard Deviation	A measure of the dispersion of the data in a frequency distribution.
Statistical Correlation	A mutual and reciprocal relationship between two or more data elements.
STF	US Department of Census Summary Tape Files. The STF files are a commonly-used source of demographic information for EJ analyses.
STORET	USEPA Storage and Retrieval of Water-Related Data System
Summary Tape File	Detailed files containing data from the U.S. Census.
Summary Statistic	One of a variety of transformations used to characterize a data set (i.e., mean, standard deviation, etc.).
Superfund	USEPA's uncontrolled hazardous waste site program created by the Comprehensive Environmental Response and Liability Act.

Surrogate Measure A demographic or environmental factor assumed to be representative of a second factor for which data are unavailable or less reliable (i.e., house value is often used as a surrogate measure for income).

TIGER/line file TIGER is the acronym for the digital (computer-readable) geographic database that automates the mapping and related geographic activities required to support the Census Bureau's census and survey programs. The Topologically Integrated Geographic Encoding and Referencing data format is commonly used in GIS analyses.

Title VI Title VI of the 1964 Civil Rights Act, 42 U.S.C. 2000(d) et.seq., as amended.

Title VI Interim Guidance EPA's Interim Guidance For Investigating Title VI Administrative Complaints Challenging Permits," issued by EPA's Office of Civil Rights on February 5, 1998, defining key parameters for the Agency's processing of Title VI administrative complaints.

Toxic Release Inventory The USEPA program which requires large-quantity generators of hazardous materials to report the nature and quantity of their annual emissions into the environment.

Tract An expanse or area of land defined utilized in demographic studies by the U.S. Census Bureau.

Transient A person who is not expected to remain in a given location for an extended period of time.

TRI USEPA Toxic Release Inventory

TRIS USEPA Toxic Release Inventory System. The TRIS database is a major source of contaminant release information for EJ analyses.

Uncontrolled Hazardous Waste Site A waste site in from which contaminants may migrate freely into the environment.

Undocumented	A citizen of a foreign country living in the U.S. without authorization of U.S. government immigration authorities.
Variability	The natural heterogeneity within a statistical population.
Weighting Factor	A factor used in a statistical analysis to represent the relative importance of an item in a population.
Welfare Standard	An established criterion developed to protect the health, happiness, and prosperity of a person, group or organization.